An interdisciplinary Digital Twin Engine for science

Andrea Manzi (EGI Foundation)

EODC Forum 2023

interTwin is funded by Horizon Europe under grant agreement n° 101058386

interTwin General Information

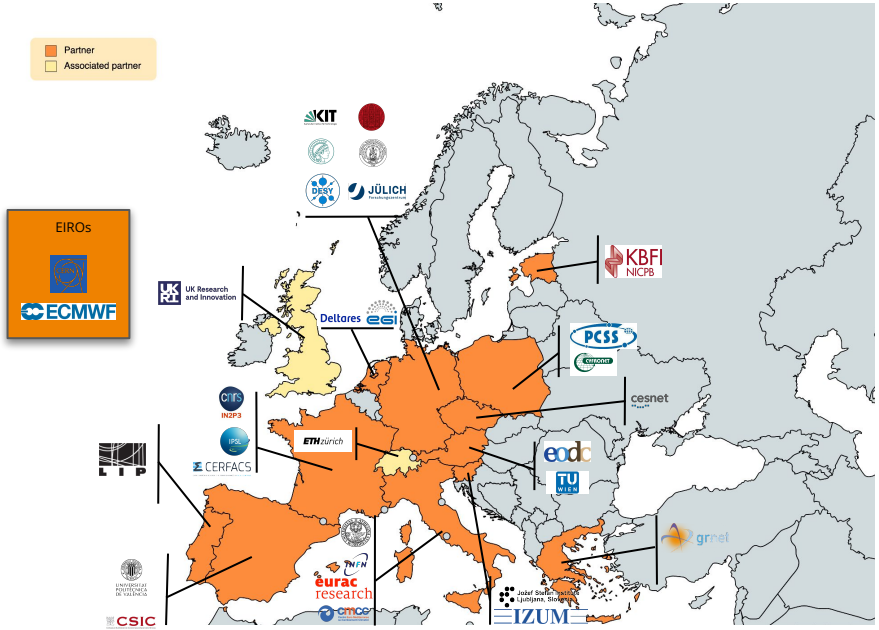
Duration	Budget
36 months	11,731,665 EUR
Period	**PMs**
1.09.22 - 31.08.25	1481.5

HORIZON-INFRA-2021-TECH-01- 01: Interdisciplinary digital twins

Outcome

- prototype of an **interdisciplinary** Digital Twin, using a combination of the latest digital technologies, to address complex challenges;
- support **interoperability** of data and software, integration and collaboration across different scientific domains;
- A framework enabling Researchers to ensure the quality, reliability, verifiability of the data available through the **Common European Data Spaces** and the **European Open Science Cloud**

interTwin Consortium Overview



EGI Foundation as coordinator

29 **Participants**, including 1 affiliated entity and 2 associated partners

Consortium at a glance

10 Providers	11 Technology providers	14 Community representants
cloud, HTC , HPC resources and access to Quantum systems	delivering the DTE infrastructure and horizontal capabilities	from 5 scientific areas; requirements and developing DT applications and thematic modules

interTwin overall objective

Co-design and implement the prototype of an interdisciplinary Digital Twin Engine.

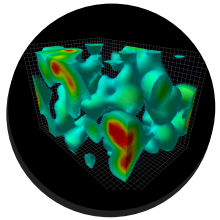
Digital Twin Engine

- It is an **open-source platform** based on open standards.
- It offers the capability to integrate with **application-specific Digital Twins**.
- Its functional specifications and implementation are based on
  - a **co-designed interoperability framework**
  - conceptual model of a DT for research - **the DTE blueprint architecture**.

Physics domain DTs

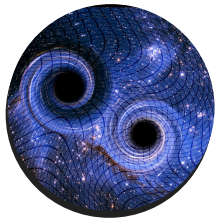
See Backup

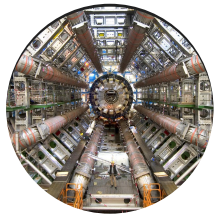
Radio Astronomy



Quantum Field Theory

Gravitational Wave Astronomy





High Energy Physics

Climate change predictions, impact and early warning for extremes events DTs

DT	Geographical region of Interest
Tropical Storms change in response to climate change	Indian and Pacific Ocean
Wildfires risk assessment in response to climate change	Europe
Flood Early Warning in coastal and inland regions	Selected European regions, Philippines
Alpine droughts early warning	European Alps
Extreme Rainfall events change in response to climate Change	Europe
Flood Climate impact in coastal and inland regions	Selected European regions, Mozambique

CERFACS
CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

IPSL

cmcc
Centro Euro-Mediterraneo sui Cambiamenti Climatici

UNIVERSITAS STUDIORUM ATHESINA

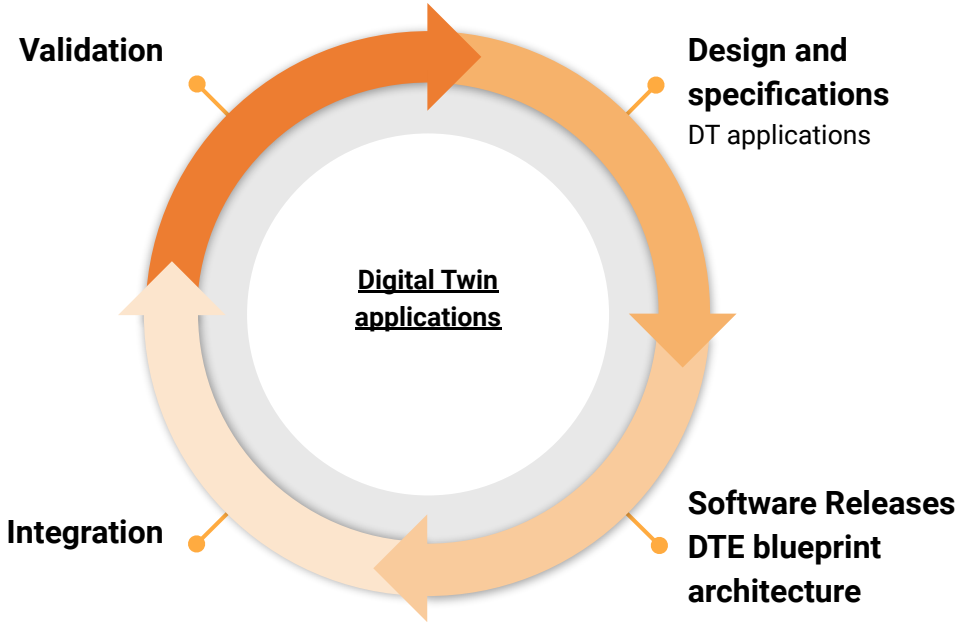
eurac research

TU WIEN

Deltares

See Next presentations

DTE Development Cycle



Final Result
Pre-operational software of a DTE

Roadmap

Project Year 1

05.2023 — **DTE Design and specifications**

Deliverables: Report on requirements for all use cases

06.2023 — **DTE blueprint architecture**

Deliverable: DTE blueprint architecture, functional specifications and requirements analysis v1

Project Year 2

10-12.2023 — **1st Software Releases**

Deliverables: Software releases for all use cases and modules

01.2024 — **DTE blueprint architecture**

Deliverable: DTE blueprint architecture, functional specifications and requirements analysis v2

04.2024 — **Validation**

Deliverables: First version of the DT applications

09.2024 — **Design and specifications**

Deliverables: Updated report on requirements for all use cases

Project Year 3

10.2024 — **DTE blueprint architecture**

Deliverable: DTE blueprint architecture, functional specifications and requirements analysis v3

01-04.2025 — **2nd Software Releases**

Deliverables: Final Software releases for all use cases and modules

07-08.2025 — **Validation**

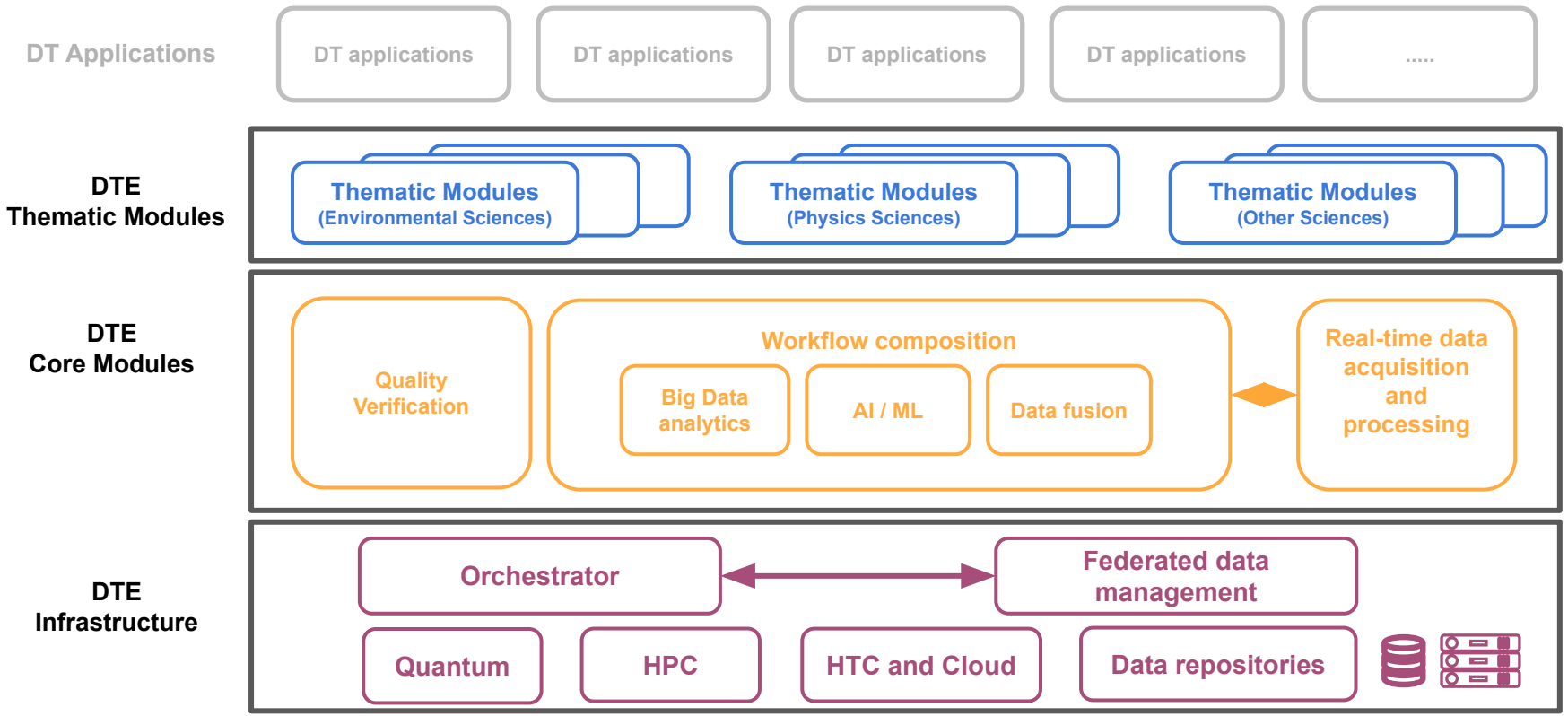
Deliverables:
DT application development and integration report
Report on software architecture concepts based on DestinE and InterTwin
Final Architecture design of the DTs capabilities

2023

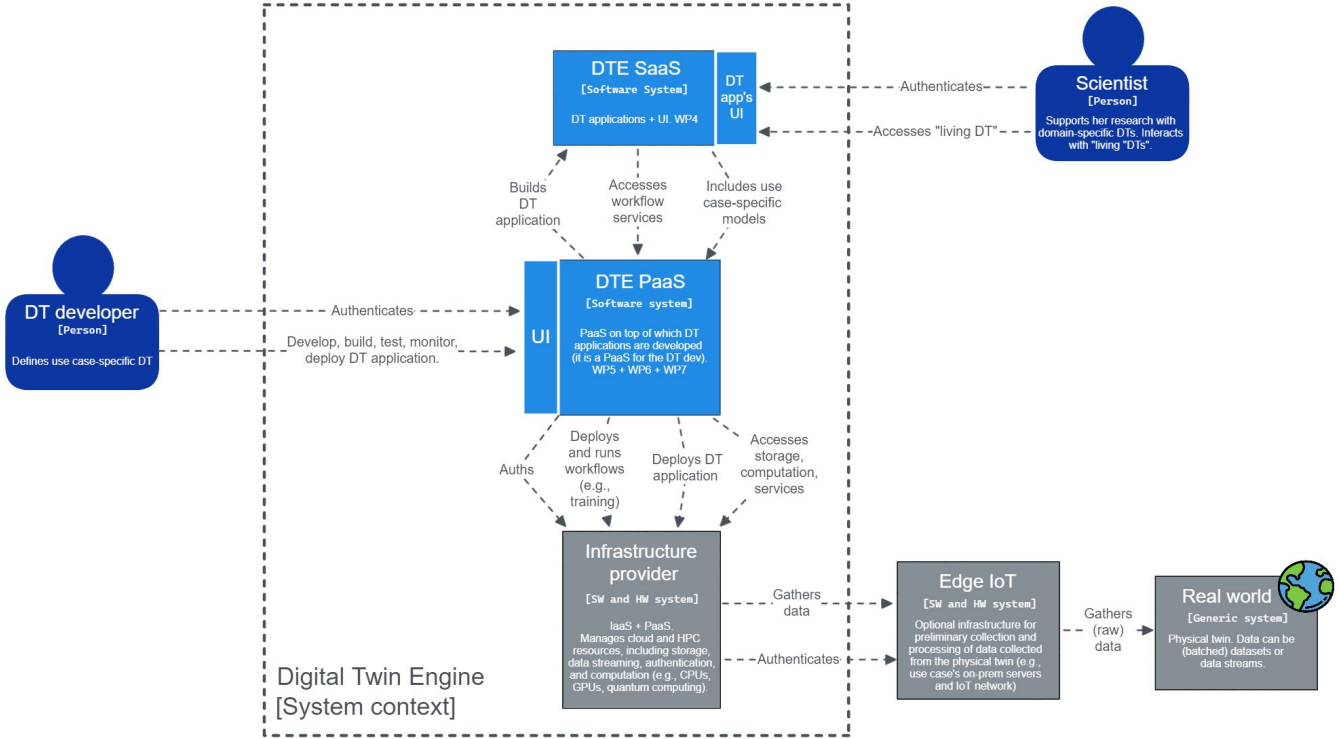
2024

2025

interTwin components



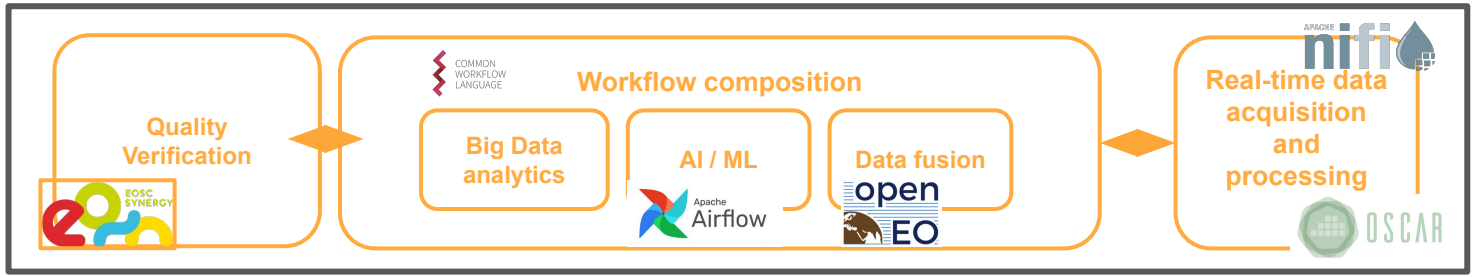
Digital Twin Engine



DTE components - DTE Core

Workflow composition	Workflow definitions based on **CWL,** run on a workflow execution system (e.g. **StreamFlow**, **AirFlow**) and able to combine self-contained execution steps from other workflow engines (e.g . **ecFlow, Ophidia, Delft-FEWS**), different back-ends, distributed big data analysis tools (e.g. **openEO, Dask, Spark**) and ML/DL training platforms (e.g., **Horovod**, **HeAT**, **PyTorch DDP**). Data Fusion as one of the workflow steps to merge observational and modeled data and different data sources.
Real-time data acquisition and processing	Generic framework for real time acquisition and processing that builds on event-triggered execution of workflow engines and exploit serverless computing. Based on **Apache NIFI** and **OSCAR** Framework
Quality Verification	Specific module for quality assurance (QA) that aims at tackling the early validation of the DTs, before being deployed as a “living DT”. Based on the **SQaaS** developed in **EOSC Synergy** project

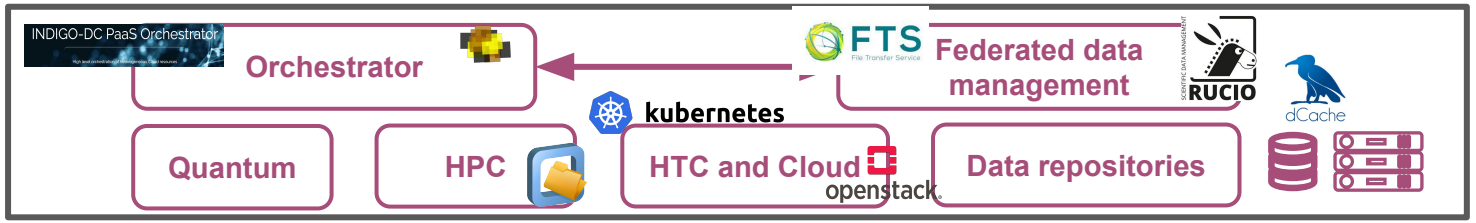
DTE Core Modules



DTE components - DTE Infrastructure

Orchestrator	**PaaS Orchestrator** + **Infrastructure Manager** elaborating deployment requested expressed in **TOSCA** being extended to deal with HPC and AI based orchestration
Federated data management	Based on **ESCAPE Data Lake** architecture and services, **Rucio**, **FTS** and HTTP accessed caches/storages. Data lake concept extended to HPC facilities
Federated Compute	Use of single-sign-on in complex simulation and modelling tasks to access data and different compute facilities, including **transparent offloading** to HPC. Automated modelling and simulation fused with data repositories and computation with containers on HTC, Cloud and HPC

DTE Infrastructure



Infrastructure providers

HPC (6)
TU Wien
GRNET
PSNC
UKRI
JSI/IZUM (EURO HPC)
JUELICH

Cloud (6+)
TU Wien
EODC
GRNET
PSNC
UKRI
JUELICH
EGI Federation

HTC (2+)
UKRI
KBFI
EGI Federation

Alignment with DestinE

- **ECMWF** is a member of the project, Task leader on interoperability with DestinE
    - Demonstrators of **data handling across interTwin and DestinE DTs** for the Extremes and Climate in production-type configurations.
    - Possible technology exchange in areas like Workflow Management and Cloud orchestration
- Had meetings with **EUMETSAT** to discuss details of the architecture of the DestinE **Data Lake** and trying to understand integrations and liaising activities
    - e.g. **openEO** is one of the main technologies we are using in the project. The main contributors are in the interTwin consortium (EODC, EURAC, WWU) and will be used as well in DestinE DL
- DG-Connect driven activity to link DestinE with interTwin and other projects (**DT-Geo**, **BioDT**, **EditoInfra**)

Conclusions

- **Challenging project**
  The DTE needs to support interoperability of data and software, integration and collaboration across different scientific domains

- **Link with EOSC**
  Aim at extending the technical capabilities of the European Open Science Cloud with modelling & simulation tools integrated with its compute platform.

- **Hybrid infrastructure**
  DTE shall enable homogeneous security and access policies, resource accounting to HPC, HTC and cloud providers

- **Bridge difference in infrastructure needs**
  The DTE shall be usable by sciences with vast differences in compute/storage needs

Thank you!

www.intertwin.eu

info@intertwin.eu

intertwin_eu

intertwin

Backup

DT Application: High Energy Physics

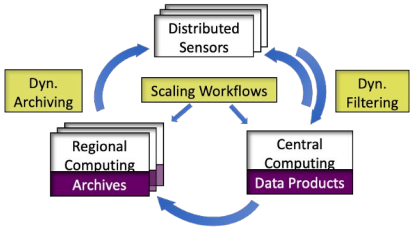
DT of Large Hadron Collider (LHC) detector components

seeking for strategies to face the increase in the need for simulated data expected during the future High Luminosity LHC runs. The primary goal is to provide a fast simulation solution to complement the Monte Carlo approach. ***Faster and deeper cycles of optimisation of the experiment parameters*** in turn will enable breakthroughs in experimental design.

DT of the Standard Model in particle physics

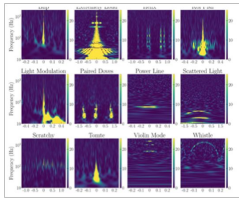
competitive results in Lattice QCD require the ***efficient handling of Petabytes of data***, therefore the implementation of advanced data management tools is mandatory. On the side of algorithmic advancement, ML algorithms have recently started to be applied in Lattice QCD. The goal is to ***systematize the inclusion of ML for large scale parallel simulations***.

DT Application: Radio astronomy and Gravitational-wave astrophysics



DT for noise simulation of next-generation radio telescopes

Providing DTs to simulate the noise background of radio telescopes (**MeerKat**) will support the identification of rare astrophysical signals in (near-)real time. The result will contribute to a realisation of "**dynamic filtering**" (i.e. steering the control system of telescopes/sensors in real-time).



DT of the Virgo Interferometer

meant to **realistically simulate** the noise in the detector, in order to study how it reacts to external disturbances and, in the perspective of the **Einstein Telescope**, to be able to detect noise "glitches" in **quasi-real time**, which is currently not possible. This will allow sending out **more reliable triggers** to observatories for multi-messenger astronomy.